DEREK THOMPSON

This Is How the AI Bubble Will Pop

The AI infrastructure boom is the most important economic story in the world. But the numbers just don't add up.

DEREK THOMPSON

OCT 02, 2025 • PAID

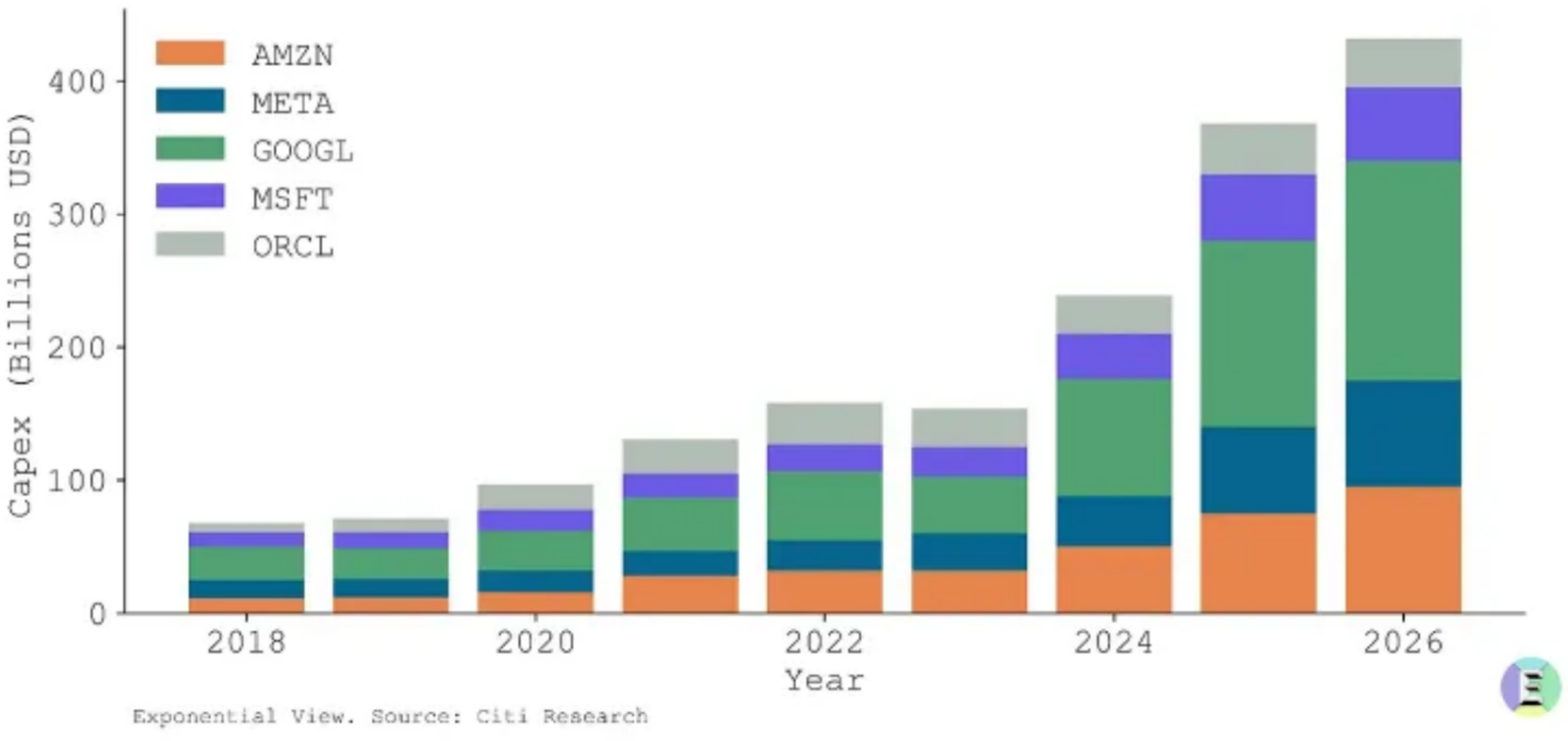
Photo by Alexas_Fotos on Unsplash

Some people think artificial intelligence will be the most important technology of the 21st century. Others insist that it is an obvious economic bubble. I believe both sides are right. Like the 19th century railroads and the 20th century broadband Internet build-out, AI will rise first, crash second, and eventually change the world.

The numbers just don’t make sense. Tech companies are projected to spend about $400 billion this year on infrastructure to train and operate AI models. By nominal dollar sums, that is more than any group of firms has ever spent to do just about anything. The Apollo program allocated about $300 billion in inflation-adjusted dollars to get America to the moon between the early 1960s and the early 1970s. The AI buildout requires companies to collectively fund a new Apollo program, not every 10 years, but *every 10 months.*

Hyperscalers' annual capex has more than doubled since ChatGPT's release

Hyperscaler annual capex (2025 and 2026 reflect estimates)



Exponential View

It's not clear that firms are prepared to earn back the investment, and yet by their own testimony, they're just going to keep spending, anyway. Total AI capital expenditures in the U.S. are projected to exceed $500 billion in 2026 and 2027—roughly the annual GDP of Singapore. But the *Wall Street Journal* has reported that American consumers spend only $12 billion a year on AI services. That's roughly the GDP of Somalia. If you can grok the economic difference between Singapore and Somalia, you get a sense of the economic chasm between vision and reality in AI-Land. Some reports indicate that AI usage is actually declining at large companies that are still trying to figure out how large language models can save them money.

Every financial bubble has moments where, looking back, one thinks: *How did any sentient person miss the signs?* Today's omens abound. Thinking Machines, an AI startup helmed by former Open AI executive Mira Murati, just raised the largest seed round in history: $2 billion in funding at a $10 billion valuation. The company has not released a product and has refused to tell investors what they're even trying to build. "It was the most absurd pitch meeting," one investor who met with Murati said. "She was like, 'So we're doing an AI company with the best AI people, but we can't answer any questions." Meanwhile, a recent analysis of stock market trends found that none of the typical rules for sensible investing can explain what's going on with stock prices right now. Whereas equity prices have historically followed earnings fundamentals, today's market is driven overwhelmingly by momentum, as retail investors pile into meme stocks and AI companies because they think everybody else is piling into meme stocks and AI companies.

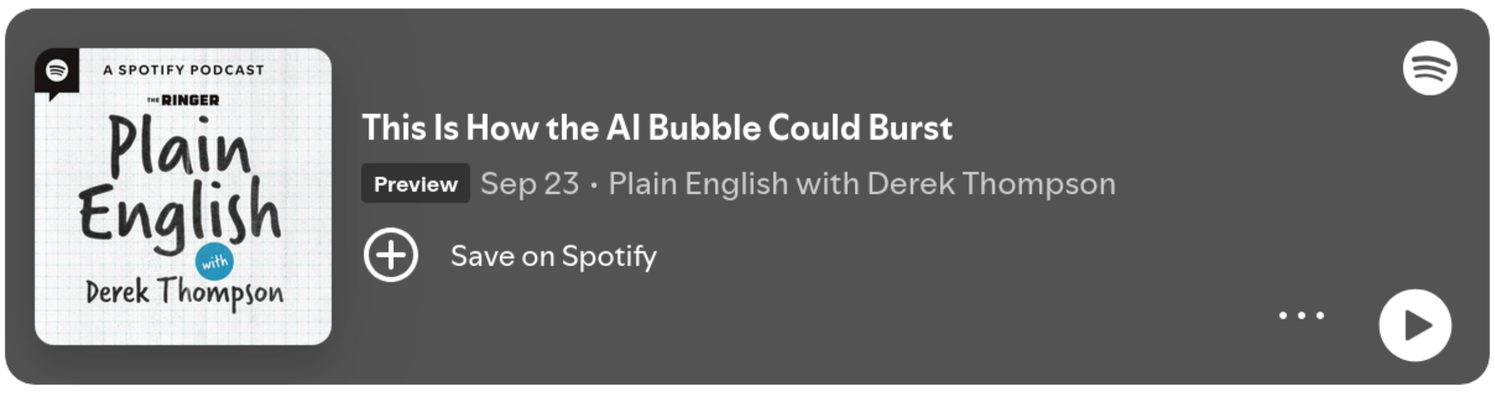
Every economic bubble also has tell-tale signs of financial over-engineering, like the collateralized debt obligations and subprime mortgage-backed securities that blew up during the mid-2000s housing bubble. Ominously, AI appears to be entering its own phase of financial wizardry. As the *Economist* has pointed out, the AI hyperscalers—that is, the largest spenders on AI—are using accounting tricks to depress their reported infrastructure spending, which has the effect of inflating their profits [1]. As the investor and

author Paul Kedrosky told me on my podcast *Plain English*, the big AI firms are also shifting huge amounts of AI spending off their books into SPVs, or special purpose vehicles, that disguise the cost of the AI build-out.

My interview with Kedrosky received the most enthusiastic and complimentary feedback of any show I've done in a while. His level of insight-per-minute was off the charts, touching on:

- **How AI capital expenditures break down**
- **Why the AI build-out is different from past infrastructure projects, like the railroad and dot-com build-outs**
- **How AI spending is creating a black hole of capital that's sucking resources away from other parts of the economy**
- **How ordinary investors might be able to sense the popping of the bubble just before it happens**
- **Why the entire financial system is balancing on big chip-makers like Nvidia**
- **If the bubble pops, what surprising industries will face a reckoning**

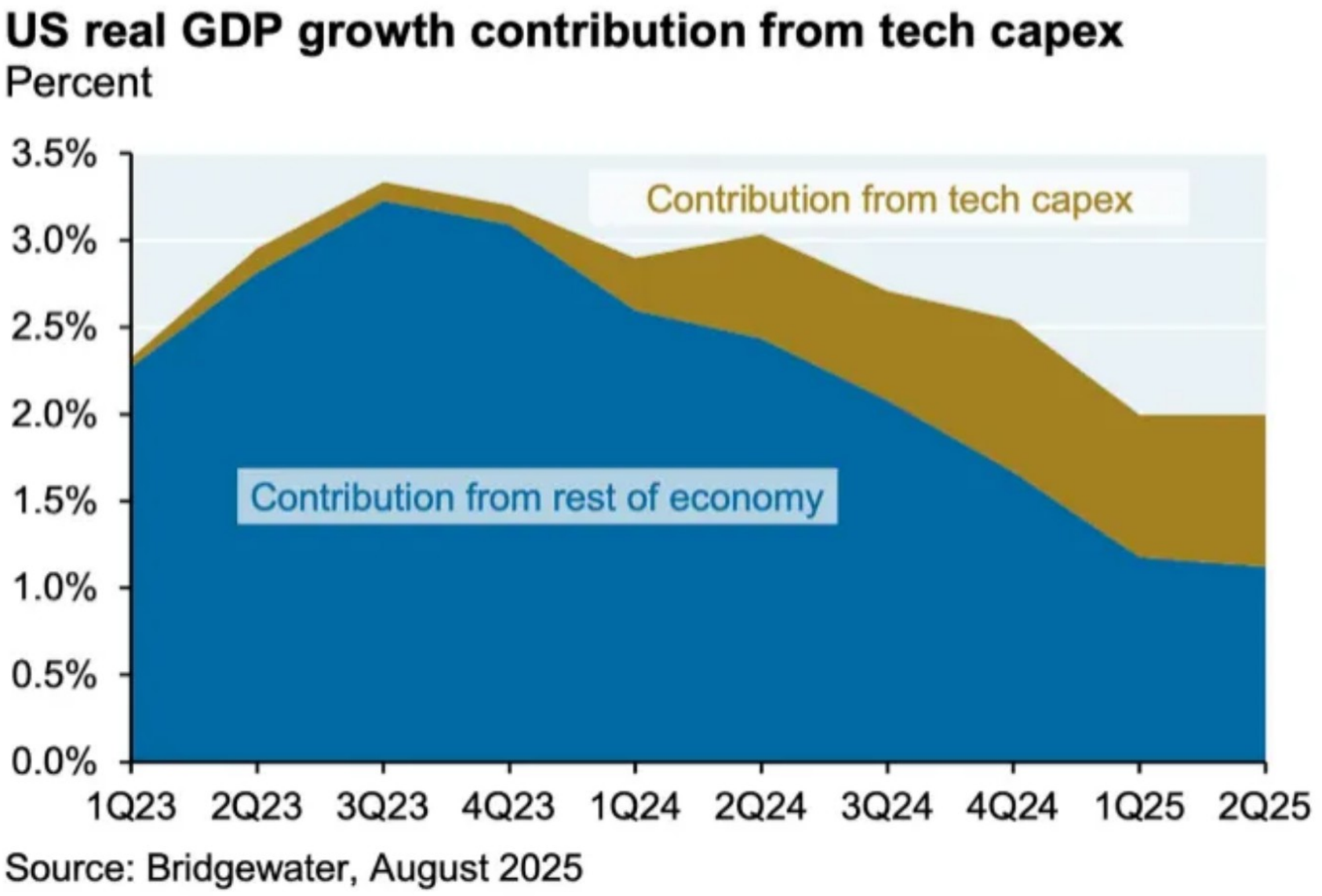
Below is a polished transcript of our conversation, organized by topic area and adorned with charts and graphs to visualize his points. I hope you learn as much from his commentary as much as I did. From a sheer economic perspective, I don't think there's a more important story in the world.



AI SPENDING: 101

Derek Thompson: How big is the AI infrastructure build-out?

Paul Kedrosky: There's a huge amount of money being deployed and it's going to a very narrow set of recipients and some really small geographies, like Northern Virginia. So it's an incredibly concentrated pool of capital that's also large enough to affect GDP. I did the math and found out that in the first half of this year, the data-center related spending—these giant buildings full of GPUs [graphical processing units] and racks and servers that are used by the large AI firms to generate responses and train models—probably accounted for half of GDP growth in the first half of the year. Which is absolutely bananas. This spending is huge.



JP Morgan chart showing the rising contribution to GDP growth from tech capex

Thompson: Where is all this money going?

Kedrosky: For the biggest companies—Meta and Google and Amazon—a little more than half the cost of a data center is the GPU chips that are going in. About 60 percent. The rest is a combination of cooling and energy. And then a relatively small component is the actual construction of the data center: the frame of the building, the concrete pad, the real estate.

HOW AI IS ALREADY WARPING THE 2025 ECONOMY

Thompson: How do you see AI spending already warping the 2025 economy?

Kedrosky: Looking back, the analogy I draw is this: massive capital spending in one narrow slice of the economy during the 1990s caused a diversion of capital away from manufacturing in the United States. This starved small manufacturers of capital and made it difficult for them to raise money cheaply. Their cost of capital increased, meaning their margins had to be higher. During that time, China had entered the World Trade Organization and tariffs were dropping. We've made it very difficult for domestic manufacturers to compete against China, in large part because of the rising cost of capital. It all got sucked into this "death star" of telecom.

So in a weird way, we can trace some of the loss of manufacturing jobs in the 1990s to what happened in telecom because it was the great sucking sound that sucked all the capital out of everywhere else in the

economy.

The exact same thing is happening now. If I'm a large private equity firm, there is no reward for spending money anywhere else but in data centers. So it's the same phenomenon. If I'm a small manufacturer and I'm hoping to benefit from the on-shoring of manufacturing as a result of tariffs, I go out trying to raise money with that as my thesis. The hurdle rate just got a lot higher, meaning that I have to generate much higher returns because they're comparing me to this other part of the economy that will accept giant amounts of money. And it looks like the returns are going to be tremendous because look at what's happening in AI and the massive uptake of OpenAI. So I end up inadvertently starving a huge slice of the economy yet again, much like what we did in the 1990s.

Thompson: That's so interesting. The story I'm used to telling about manufacturing is that China took our jobs. "The China shock," as economists like David Autor call it, essentially took manufacturing to China and production in Shenzhen replaced production in Ohio, and that's what hollowed out the Rust Belt. You're adding that telecom absorbed the capital.

And now you fast-forward to the 2020s. Trump is trying to reverse the China shock with the tariffs. But we're recreating the capital shock with AI as the new telecom, the new death star that's taking capital that might at the margin go to manufacturing.

Kedrosky: It's even more insidious than that. Let's say you're Derek's Giant Private Equity Firm and you control $500 billion. You do not want to allocate that money one $5 million check at a time to a bunch of manufacturers. All I see is a nightmare of having to keep track of all of these little companies doing who knows what.

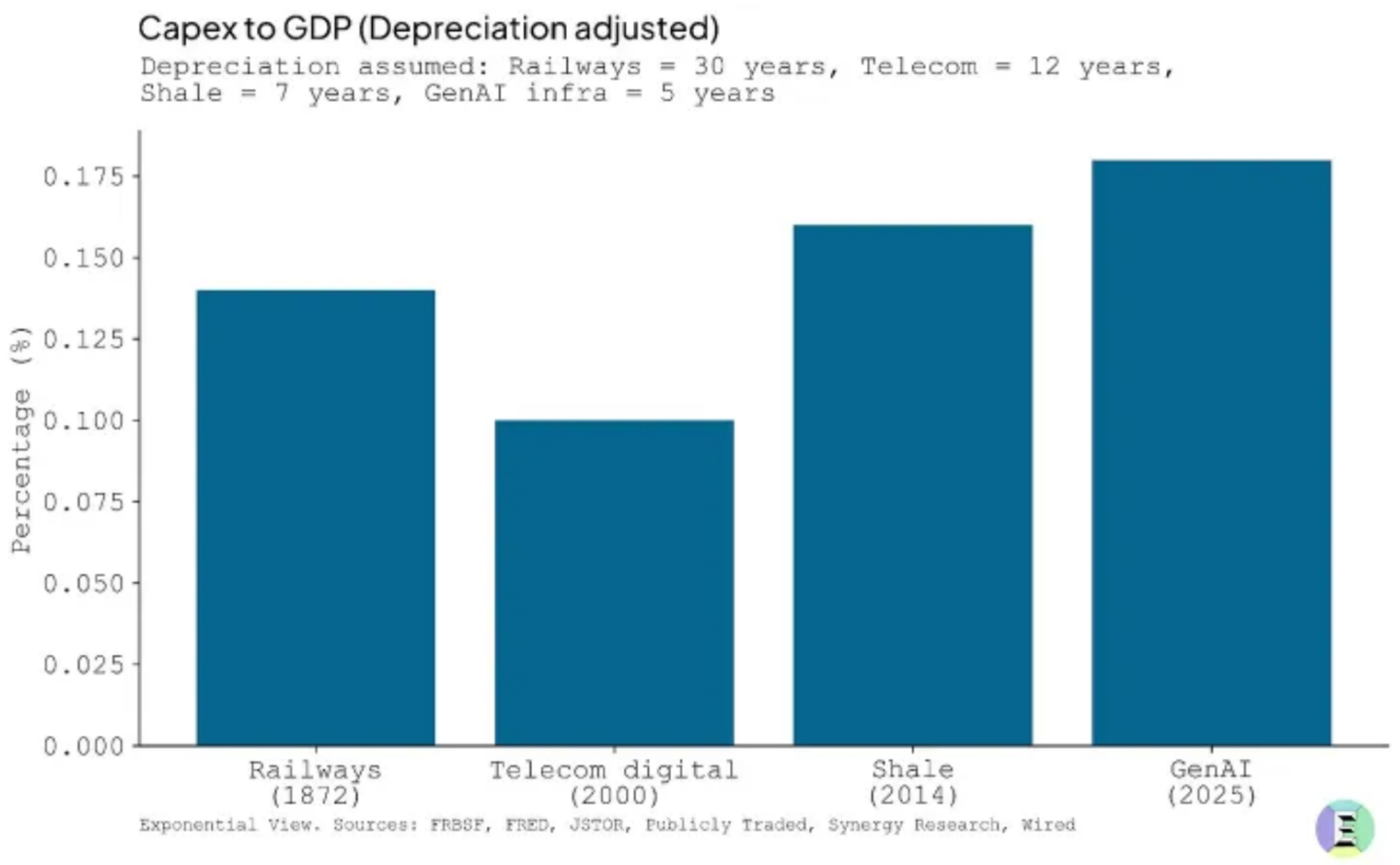
What I'd like to do is to write 30 separate $50 billion checks. I'd like to write a small number of huge checks. And this is a dynamic in private equity that people don't understand. Capital can be allocated in lots of different ways, but the partners at these firms do not want to write a bunch of small checks to a bunch of small manufacturers, even if the hurdle rate is competitive. I'm a human, I don't want to sit on 40 boards. And so you have this other perverse dynamic that even if everything else is equal, it's not equal. So we've put manufacturers who might otherwise benefit from the onshoring phenomenon at an even worse position in part because of the internal dynamics of capital.

Thompson: What about the energy piece of this? Electricity prices rising. Data centers are incredibly energy thirsty. I think consumers will revolt against the construction of local data centers, but the data centers have enormous political power of their own. How is this going to play out?

Kedrosky: So I think you're going to rapidly see an offshoring of data

centers. That will be the response. It'll increasingly be that it's happening in India, it's happening in the Middle East, where massive allocations are being made to new data centers. It's happening all over the world. The focus will be to move offshore for exactly this reason. Bloomberg had a great story the other day about an exurb in Northern Virginia that's essentially surrounded now by data centers. This was previously a rural area and everything around them, all the farms sold out, and people in this area were like, wait a minute, who do I sue? I never signed up for this. This is the beginnings of the NIMBY phenomenon because it's become visceral and emotional for people. It's not just about prices. It's also about: If you've got a six acre building beside you that's making noise all the time, that is not what you signed up for.

A VERY SPECIFIC PREDICTION FOR HOW AND WHY AI BUBBLE WILL POP ...



Exponential View

Thompson: You've compared the infrastructure boom to the railroads and the dot-com fiber-optic build-out. The rail we laid in the 1860s was still around in the 1890s. The fiber-optic cable worked well for years. But the most advanced GPUs are getting updated all the time. So this isn't quite like building a railroad. At the opposite extreme of railroads, there are bananas. If you spend $300 billion on bananas today, your capex isn't worth anything in two weeks, because all the bananas are brown. I don't think GPUs are bananas. But they're also not steel rails.

Kedrosky: They're closer to bananas than steel. [2]

Thompson: What does this tell us?

Kedrosky: The market is rewarding [the big tech companies] for investing in AI even though it makes no economic sense to spend at this level because there's no way they can recoup the value of the capital spending over the next three years. So they'll be forced to do these kind of wacky shell games where they say, "Well, the building itself will actually be valuable in five years, because it'll still have energy, it'll still have water, it'll still be able to cool things, the walls will still be standing, and I'll just swap out the GPUs." But the problem is *the GPUs are the majority of the cost*. The shell is the thing I'd like to write off, since I don't want to have to write off GPUs every three years. But they're the majority of the cost of what we call a data center.

Unlike telecom, unlike the fiber boom, unlike in railroads, there are actually two assets here. One that's long-lived, a building, which is essentially a small fraction of the cost of the center; and one that's very short-lived, which is the GPUs, which are the thing we'd like to have last and don't, yet represent as much as 60 percent of the cost of the data center. So there's the perversity.

Thompson: I want to talk about how some of this might go badly in the next few years, and I want to preface that discussion by saying that when I talk about AI as a bubble, I think some people see me as being pessimistic about the technology. The railroads were a bubble. There was a panic of 1857, of 1873, and of 1893. There were constant railroad depressions, and also the railroads changed the world. Broadband was a bubble, it also changed the world. Big infrastructure buildouts that changed the world often passed through a bubble phase. So it's not *pessimistic* to say that AI is currently in a bubble. You could say it's actually historically in tune to say that we are very likely in the middle of a bubble, because every industrial revolution passes through bubble phases.

Let's start here. How close are the hyperscalers—Meta, Google, Microsoft, the big boys—to getting AI revenue to match AI spending?

Kedrosky: Nowhere near.

The hyperscalers are spending as much as 50 percent of income on capital expenditures, which is unprecedented. This doesn't happen. Normally, if I did that as Microsoft or Amazon, I would be taken to the woodshed and beaten by investors because that's such an incredible investment on one narrow slice of CapEx. They're not being punished for that.

What I'm watching is how they're moving the financing off their balance sheet. That for me is a reflection of not wanting the credit rating agencies to look at what they're spending. What we're seeing is these SPVs— special purpose vehicles—being created. Meta has a stake, some giant private debt provider has a stake, and the data center at the end is under Meta's control, but they don't "own" it. And so it doesn't go on their balance sheet in terms of assessing

creditworthiness. We're seeing for the first time over the last six, seven months, the beginnings of a wave of these special purpose vehicles and other more exotic financing structures. We're seeing the equivalence of some of the old collateralized debt obligations emerge. These are all, for me, the beginning of the sign that the bubble is becoming tired because the market is beginning to punish—at least there's a perception that the market will punish—if I continue to keep this on my income statement. So I move it somewhere else. And that makes the entire process much more opaque. That's the thing to watch. How hard are they trying to hide the expenditure?

Thompson: I love the way you put it. It's kind of like "you know your behavior is unethical if you try to keep it a secret." You know your economic activity is bubblicious if you try to dress it up in financial opacity.

Let's talk about just exactly how this works. The big guys, the hyperscalers, are getting together with private capital and they're putting money into this box called a special purpose vehicle. The box is the thing that's investing in these data centers. Take me through exactly how this works.

Kedrosky: The private credit firms, like Apollo, want a stake in data centers. From their standpoint, I can write a large check and if I partner with Meta or Google, because there's a high likelihood that immediately people will be using the data center. It'll be populated, and there'll be rental income flowing back from it.

So now I have multiple people participating, in part because their interests are aligned, but also from the standpoint of these individual public companies, because I don't have to roll it up into my income statement if I control less than 50 percent of it. From a legal standpoint, I don't want legal control because then that flows back and I have to deal with it from the standpoint of my credit rating, in terms of the leverage, the amount of debt I have on my balance sheet against my equity, and all of these things that are really mundane and boring, but matter immensely to CFOs.

So from their standpoint, the idea of partnering with a large private credit firm to create these bespoke special purpose vehicles that are these one-off vehicles, they're great because now I get what I wanted, which is a new data center, and I don't get what I didn't want, which is a hit to my credit rating. And so there's a huge incentive to create these.

Thompson: Let me try to restate this so I understand it. So Meta wants to build these gigantic AI data centers. These projects cost tens of billions of dollars altogether. Meta is rich, but they don't want to just borrow all the money the normal way. They don't want the spending to show up on their balance sheets. So they solve the problem by creating this special box. Meta puts some assets into the box. Another private

investor puts some money in the box. And now that box, a special purpose vehicle, is going out and borrowing money, paying for construction, and owning the data center. On the surface, everyone's happy. Meta gets money without messing up its balance sheet. The investors get high returns without high risk because they're basically working with Meta.

Now what happens if we build too many data centers? Meta is exposed. The private capital firms are exposed. What happens?

Kedrosky: Not to go full hedge fund, but if you think about which companies really benefit, they're in construction and air conditioning. Carrier, for example. If I'm building out industrial-class air conditioning for giant data centers at an unprecedented scale, what business would I really want to be in? Selling air conditioning. Cooling those buildings is a huge problem. If I'm an industrial provider of air conditioning, this is a great time. These companies aren't obvious, but they're huge beneficiaries of the build-out.

Increasingly, investments in special-purpose vehicles and other data-center spending are showing up in REITs—real estate investment trusts. If you look inside any large REIT in the U.S. today, somewhere between 10 and 22 percent is already directly data-center related. So if you're a conservative investor with a REIT in your portfolio—because you think, "I don't care about that crazy tech stuff, I'll just be safe in commercial real estate, collecting income"—go look inside your REIT. Two years ago, there was nothing related to data centers. In some of the largest ones today, 21 or 22 percent is directly tied to data centers. You're already in it, my friend.

Thompson: You just told me that data center costs are like 60 percent chips. That means these REITs are significantly in Nvidia stock, right now? It just seems like this whole economy balances on the narrow reed of Nvidia.

Kedrosky: It's not funny, but it is kind of funny. Imagine I'm a risk-averse investor, and I say, "You know what, I'm only in index funds." A year or two ago, someone might have told you, "Well, you may think you're being cautious, but 30 percent of the S&P 500 is now tied to what's euphemistically called the Magnificent Seven stocks. You're actually long Nvidia in a huge way."

So you decide, "I'm getting out of the S&P 500 and going into really safe stuff, like REITs." But it's the same problem. There's nowhere to run. Increasingly, you've got nowhere to hide. In a backdoor way, private credit is now allowed inside retirement funds, and you're seeing it show up in other places—not just in REITs, but even in private credit ETFs. You think you're investing in something safe, like a take-private operation from a manufacturer in Iowa. No, you're not. You're in data centers. And by proxy, you're also in Nvidia.

So what looks like a complex system actually has a single point of failure: a handful of semiconductor stocks, highly leveraged to everything that's happening. They've metastasized across the entire system—from the S&P 500 to REITs to private credit to ETFs. It's incredibly insidious and important, and most people don't realize how deeply it's insinuated itself.

Thompson: What does a bubble look like to you? What are the news headlines?

Kedrosky: For starters, the largest share of future data-center building is happening through SPVs. When you see people saying, "Look, it's all being done in partnerships—it's not as risky for Meta or Amazon because they're partnering"—that's a hallmark of a bubble. The companies themselves are stepping away aggressively because they see the risks.

Another sign: delays in air conditioning and other critical equipment. Interconnect gear for racks and GPUs was delayed four or five months at one point. If those delays suddenly vanish, that's actually bad news—it means demand is slowing. So if you hear about industrial-class air conditioners missing their numbers, that's a red flag. Right now, the only reason they'd miss is because data centers aren't buying.

These are the things to watch—the edges—not anecdotes about someone trying OpenAI and not liking it. The bigger risk is the capital being sucked out of small manufacturers who might otherwise be onshoring and hiring, but now can't compete. That's how you end up making bad policy decisions: looking at GDP growth of 3 percent and saying, "See, the economy's fine," without recognizing how much of it is tied to transient data-center spending.

I make this dumb joke all the time, but it's like my dog barking at the mailman. He thinks barking makes the mailman go away. We're that barking dog—confusing causality, thinking growth comes from tariffs or something else, when it's really being driven by data centers. That's when you know you're in a bubble—when people double down on the wrong things because they've misunderstood the cause.

'I DON'T SEE HOW THIS CAN CONTINUE'

Thompson: The economic commentator Noah Smith wrote about what it would look like if a data-center bubble became a true financial crisis. He laid out four conditions:

1. A big "this time is different" story around AI.
2. A large and growing pile of debt concentrated in one sector.
3. An opaque corner of the financial system—private credit—that's

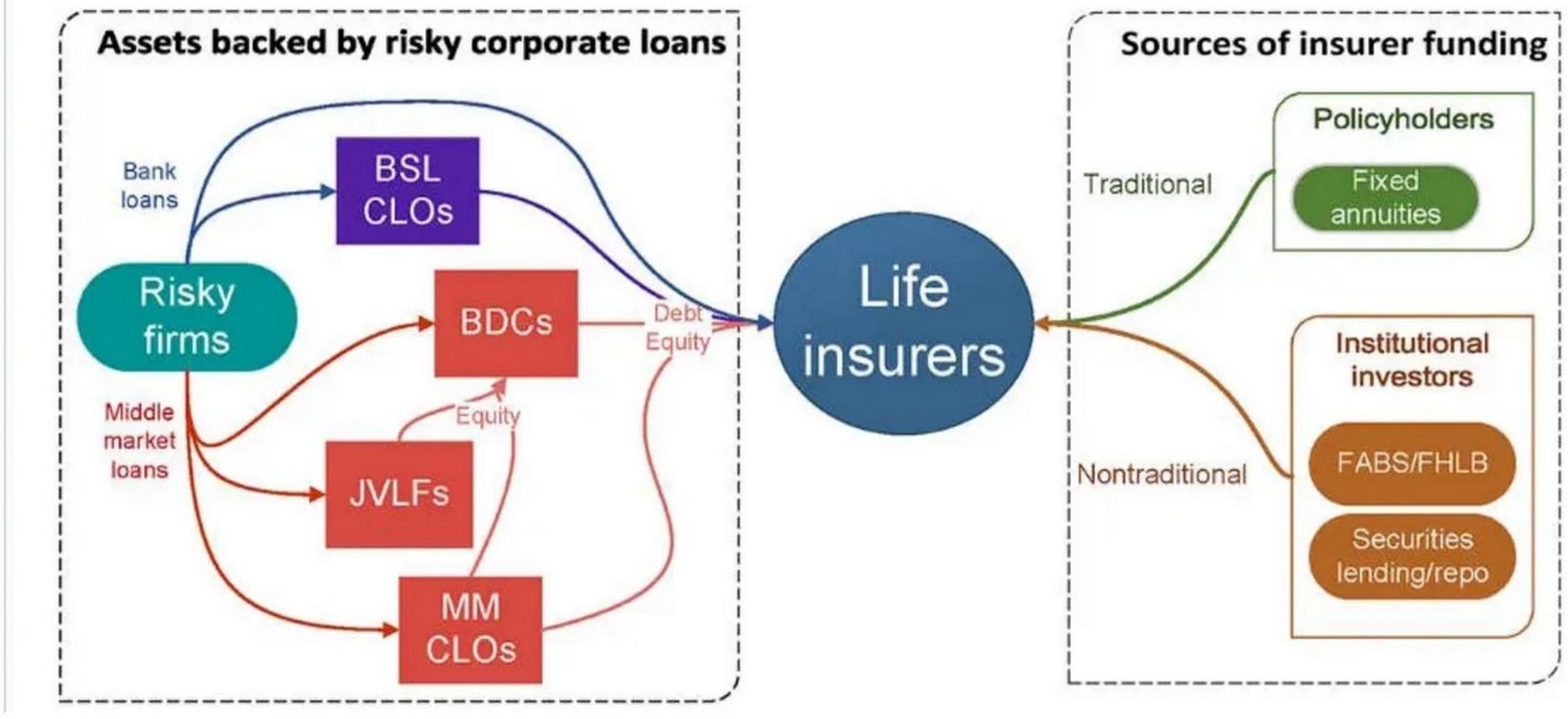
expanded rapidly.

4. Systemically important lenders, including banks and life-insurance companies, heavily exposed.

Do you see those as the ingredients of a real financial crisis?

Kedrosky: Absolutely. It has all the pieces. I'll pick on just one that you mentioned: insurance. Private equity and private credit have purchased insurance companies. Insurance premiums get reinvested in whatever the private equity firm is doing. This creates a classic timing mismatch. The data centers are relatively transient. The obligations to insurance policyholders are longer term. So I have mismatched assets and liabilities. That's no different than what happened way back during the financial crisis, which did in Bear Stearns. They had lent long and owed short, and so they ended up blowing up. That's a new, under-recognized source of systemic risk. Life insurance firms are increasingly owned by private equity firms. That's not nearly well understood enough.

Figure 1. Life insurers' position in the intermediation chain of credit to risky firms



Noah Smith: "Life insurers have basically become a kind of bank that borrows money from institutional investors and lends it to private credit."

Thompson: Based on the math that you're describing, when is it reasonable to think that some kind of break could appear in the system?

Kedrosky: Current projections would tell you that in two or two-and-a-half years out, [the tech companies] are no longer earning a risk-adjusted return commensurate with the cost of doing business as a data center. A naive projection puts you out like two and a half years. It's hard to imagine absent something changing materially with respect to the amount that companies are earning from selling AI services, that it doesn't happen even faster than that. So my guess is two to two and a half years. That's when you really have to be saying to yourself, "I don't see how this can continue."

---

1 The basic claim is: The big tech companies, such as Microsoft, Amazon,

Alphabet, and Oracle, are spending a bunch of money on advanced computer chips to train and use AI. In their financial reports, they get to decide how to report this spending based on how long it takes those chips to lose their value. This is called depreciation. If a building is assumed to last a long time, the company can reasonably spread its cost over many years, because it's still getting value from the investment across those years. By reporting smaller annual expenses, this makes their profits look bigger in the short run. But Nvidia chips are only cutting-edge for a few years—maybe two or three—because the company keeps coming up with better models. If the big tech companies are spreading their chip spending over six years despite knowing that they're going to have to buy new chips in three years, they're understating their real costs and overstating their profits.

2 The defense from AI folks here will be that less advanced chips can still run inference, or process LLM prompts, for years after they're no longer cutting edge. But if cheaper chips can run AI, it raises the question of whether today's money is being spent wastefully. An imperfect but directionally useful analogy: Imagine I'm building a house to rent out through Airbnb, and I buy tons of a very expensive lumber—say, African blackwood. "Derek, you've bought way too much wood for a house this size," you tell me. "Who cares," I say, "whatever I don't put into the house can be kindling for future guests." "Sure," you say, "but cheap pine wood makes just as good kindling." The problem, essentially, is that Nvidia's chips might be African blackwood in 2025 but they'll be pine wood by 2028.

---

595 Likes • 95 Restacks